IRIS AperTO

UNIVERSITÀ DEGLI STUDI DI TORINO

AperTO - Archivio Istituzionale Open Access dell'Università di Torino

The Questionnaire of Cognitive and Affective Empathy: A Comparison Between Paper-and-Pencil Versus Online Formats in Italian Samples

This is the author's manuscript

Original Citation:

Availability:

This version is available http://hdl.handle.net/2318/1686434 since 2019-01-10T11:43:19Z

Published version:

DOI:10.1080/00223891.2017.1389745

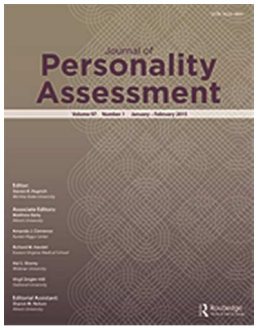
Terms of use:

Open Access

Anyone can freely access the full text of works made available as "Open Access". Works made available under a Creative Commons license can be used according to the terms and conditions of said license. Use of all other works requires consent of the right holder (author or publisher) if not exempted from copyright protection by the applicable law.

(Article begins on next page)

05 May 2023



The Questionnaire of Cognitive and Affective Empathy: A Comparison between Paper-and-Pencil versus Online Formats in Italian Samples

Journal:	*Journal of Personality Assessment*
Manuscript ID	Draft
Manuscript Type:	General Submission
Keywords:	Empathy, QCAE, Online, Italian, Well-being

For Peer Review Only

SCHOLARONE™
Manuscripts

ONLINE VS. PAPER-AND-PENCIL ITALIAN QCAE

Abstract

The most recent conceptualizations of empathy recognize affective empathy as distinct from cognitive empathy. Consequently, instruments that assess these two types of empathy have been developed. Among them, the Questionnaire of Cognitive and Affective Empathy (QCAE) is a particularly promising, relatively new, self-report measure consisting of 31-items. To examine the cross-cultural adaptability of the QCAE, we investigated the psychometric properties of an Italian version in two samples and with two different formats of administration. Study 1 (n = 407) used archival data collected via paper-and-pencil, whereas Study 2 (n = 285) administered the QCAE along with some other measures via an online format. Factor structure, internal consistency, and convergent validity were tested. The findings of both studies provide support for the cross-cultural applicability of the QCAE, and reveal interesting associations between empathy and interpersonal competence, well-being, and emotion dysregulation.

Keywords: empathy; QCAE; online format; Italian; well-being.

For Peer Review Only

The Questionnaire of Cognitive and Affective Empathy:

A Comparison between Paper-and-Pencil versus Online Formats in Italian Samples

Although scholars have attributed different meanings to the word "empathy" over time, all agree that this psychological construct plays a key role in human interactions. Empathy was first defined as "to feel into" (Lipps, 1903) and referred to how people come to know others' emotional states. Since then, many refinements of this complex and multifaceted construct have been proposed (Leiberg & Anders, 2006). The importance of empathy is self-evident; every time we try to understand others' behaviors or intentions, and we consequently adapt our social behavior to improve our interactions and relationships, we are empathizing. Without some empathetic skills, human beings would be seriously limited in building interpersonal relationships, because they would be blind to others' needs and desires.

Different definitions of empathy have been proposed in various contexts such as psychotherapy, social psychology, neuropsychology, and even ethology. Since the construct of empathy encompasses some seemingly opposing aspects – "cognitive versus affective, attitude versus behavior, a momentary experience versus a life situation, shallow versus deep, and expressed versus unexpressed" (Bohart & Greenberg, 1997, p. 444) – some authors have tried to put forward an all-encompassing definition of empathy that takes into account its whole complexity. For example, Rogers (1959) proposed that empathy would be the capability to perceive the internal frame of reference of someone else with the same emotional components and meanings. This definition is probably incomplete, when compared to more recent conceptualizations, as it only focuses on the voluntary act of "put[ting yourself] in someone else's shoes," without taking into account the emotion-contagion process, which is mainly an automatic and spontaneous one.

For Peer Review Only

Hoffman (1984) – on the contrary – focused more on the emotional aspects of empathy, and defined it as an "affective response more appropriate to someone else's situation than on one's own" (Hoffman, 1984, p.114).

In an attempt to integrate these seemingly contradictory definitions, Bateson (2009) more recently hypothesized that a good way to conceptualize empathy would be to consider it as *'the answer'* to the following two questions. First, how can a human being know what someone else is thinking or feeling? Second, what leads this person to react with sensitivity in front of his or her suffering?

With regard to the first of these questions, Preston and de Waal (2002) introduced the so-called perception-action model of empathy. The core of this theory is that the empathetic process is automatically triggered by the view of the emotional state of another person. This process starts with a corresponding representation of that emotional state, based on a somatic and a motor reaction in the observer, and then it could conduct to an empathetic behavior. This model is supported by the somatic marker hypothesis (Damasio, 1994) and the discovery of mirror neuron system (MNS; Rizzolatti, 1996), given that it is based on the idea that the experience of the observer overlaps with the experience of the observed person. However, everyday life shows that we can vicariously experience emotional states of others, or deduce them, even if we have never experienced similar situations before (e.g., war-related emotions, earthquakes, etc.). In addition, sometimes it is sufficient to hear, to read or even to imagine about others' experiences to provoke an empathetic feeling. Thus, a perception-action model is necessary but not sufficient to explain the entire empathy process.

To overcome these limitations, Baron-Cohen et al. (2005) have proposed a model of empathy that better accounts the fact that we can be empathetic without an emotional contagion-like

URL: http://mc.manuscriptcentral.com/JPersAssess Email: jpa_office@emich.edu

process. More in detail, they identified two levels of the empathy processes: the lower one, which develops early, is the affective part of empathy and includes the contagion-like process; the higher one, which develops later, is the cognitive part of empathy and includes complex cognitive processes like the Theory of Mind (Leiberg & Anders, 2006).

With regard to the second of the questions posed by Bateson (i.e., what leads a person to react with sensitivity in front of suffering), Preston and de Waal (2002) suggested that empathy is a prosocial behavior based on the cost/benefit in peer and kin groups. In this model, support, assistance, and help become advantageous because they are likely to be reciprocated by other members of the group (see, for example, Trivers, 1971). Another possible account for the humans' predisposition toward being sensitive or empathetic to others' suffering could be found in the emotional contagion process (Weisbuch, Ambady, Slepian, & Jimerson, 2011). In this view, the exposure to others' pain would automatically and intrinsically elicit some distress in the observer, and therefore the observer could choose to act empathetically simply because s/he wants to discontinue his or her own distress.

Cognitive and Affective Empathy

Despite the technical differences from one model to another, nowadays most authors agree that the construct of empathy might be broken down into two components: the first refers to the understanding of other people's emotions and the second refers to vicariously experiencing them. These two abilities have often been conceptualized under the labels "cognitive empathy" and "affective empathy" (Gibbons, 2011; Leiberg & Anders 2006; Reniers et al., 2011).

Cognitive empathy, more in detail, is the ability to understand how other people might feel, using visual, auditory, and/or situational cues. Within the research literature (e.g., Blair, 2005), cognitive empathy is sometimes considered to be a construct very close to that of Theory

For Peer Review Only

of Mind (ToM; Lawrence, 2004; see also Kanske, Böcckler, Trautwein, & Singer, 2015; Brown, Thibodeau, Pierucci, & Gilpin, 2017), which is the capability to understand that we and others have mental states, that mental states of others can differ from one's own, and that others' behavior can be explained by their mental state (Frith & Frith, 2003). It should be noted, however, that while ToM is more concerned with the acknowledgment of mental states, such as *intentions, desires,* or *beliefs* (Völlm et al., 2006), cognitive empathy is more focused on recognizing others' *emotional experiences* and *feelings* (Jolliffe & Farrington, 2004; Reniers et al., 2011).

As for affective empathy, it refers to a person's emotional reaction to other people's experiences and it does not necessarily require cognitive understanding (Leiberg & Anders, 2006). However, affective empathy is not just a contagion-like process, because the emotional responses do not necessarily match those of the target (e.g., one may feel compassion or tenderness for someone who is feeling frightened). Moreover, not all emotional responses could be considered empathic (e.g., schadenfreude, happiness about another's tragedy; Feather & Nairn, 2005), and therefore, the emotional reaction to a social stimulus should be "other-oriented" in order to be considered "empathic" (Lawrence, 2004).

Empathy and Psychopathology

The Diagnostic and Statistical Manual of Mental Disorders, Fifth Edition (DSM-5; APA, 2013) contains two personality disorders characterized by a lack of empathy: Narcissistic Personality Disorder (NPD) and Antisocial Personality Disorder (ASPD). In NPD, lack of empathy is explicitly indicated in the general definition of the disorder: "A pervasive pattern of grandiosity, need for admiration, and lack of empathy" (APA, 2013, p. 669). In ASPD, although there is not a clear reference to lack of empathy, its criteria include language highly suggestive of em-

URL: http://mc.manuscriptcentral.com/JPersAssess Email: jpa_office@emich.edu

For Peer Review Only

pathic deficits: "Lack of remorse, as indicated by being indifferent to or rationalizing having hurt, mistreated, or stolen from another" (APA, 2013, p. 659). Another example, from a different theoretical conceptualization of psychopathology, can be found in the Psychodynamic Diagnostic Manual (PDM; Force, 2006). In this manual, Sociopathic Personality Disorder or Narcissistic Personality Disorder in children and adolescents are characterized by un(der)developed empathic abilities or concern for others (Force, 2006).

Assessment of Empathy

Because of its importance to understanding and assessing mental disorders, a number of assessment instruments to measure empathy have been developed. The most widely utilized tools are the Interpersonal Reactivity Index (IRI; Davis, 1980) and the Empathy Quotient (EQ; Baron-Cohen, 2004). The IRI is a 28-item self-report measure of interpersonal competencies consisting of four scales. According to Davis (1980), the Fantasy scale and Perspective Taking scale would assess cognitive empathy. Fantasy concerns the empathic response to fictional characters, and Perspective Taking assesses the capability to assume the point-of-view of other people. Conversely, Empathic Concern and Personal Distress would assess affective empathy. Empathic Concern refers to the sympathetic response to others' feelings, Personal Distress is a measure of self-oriented feelings of anxiety (Davis, 1980). The EQ comprises 60 questions, broken down into two types: 40 questions about empathy and 20 distractor items. An unique, global, empathy score is produced by summing up all responses to the 40 questions concerning empathy.

A few years ago, based on the most recent conceptualizations empathy viewing it as a multidimensional construct, Reniers and colleagues (2011) developed the Questionnaire of Cognitive and Affective Empathy (QCAE), a measure aimed at assessing empathy along three dimensions: cognitive, affective and total. Compared to other similar, empathy measures, the

For Peer Review Only

QCAE offers the advantage of providing separate, reliable scores for the cognitive and affective components of empathy.

The Current Study

The QCAE has been introduced recently, and independent research on its reliability and validity is needed. Specifically, the QCAE has not been thoroughly investigated in non-English speaking samples, and no independent validation studies of the QCAE have been conducted yet. Furthermore, although emerging research suggests that the format with which a questionnaire is administered (in terms of paper-and-pencil vs. online) should not dramatically affect its scores (Ritter, Lorig, Laurent & Matthews, 2004; Riva, Teruzzi, & Anolli, 2003), to date no studies have yet examined whether there are any differences between a paper-and-pencil vs. an online administration format of the QCAE. This type of research is particularly important as most questionnaire studies are currently conducted using an online format (Skitka & Sargis, 2006). Therefore, the current study aimed to: (1) provide information on the reliability and validity of the Italian version of the QCAE, and (2) compare the outcomes of paper-and-pencil vs. online administration formats.

Materials and Methods

This study used two datasets derived from two research projects, which used different methods of administering the QCAE (paper-and-pencil versus online). The first project aimed at investigating an interpersonal competence measure, and used the QCAE in its paper-and-pencil version, to investigate convergent validity (Giromini et al., 2015). The second project aimed to examine an online format for QCAE administration.

Participants

For Peer Review Only

Paper-and-pencil Dataset. The paper-and-pencil dataset consists of data from a study conducted by Giromini et al. (2015). After translating the questionnaire to Italian, using the translation-back translation method, the authors administrated the QCAE in paper-and-pencil format. The original sample size of Giromini et al.'s (2015) study consisted of 408 students from an Italian university, ranging in age from 18 to 57 (M = 22.6, SD = 4.6), 74% were women (Giromini et al., 2015). However, one of the participants did not fill out the QCAE, so that our final sample was reduced to 407.

Although the authors inspected central tendency, dispersion, and internal consistency, Giromini et al. (2015) did not present detailed analyses on the reliability and validity of the Italian QCAE.

Online Dataset. The online dataset was collected to evaluate the psychometric properties of the QCAE obtained by means of online administration. The same Italian QCAE, translated into Italian by Giromini et al. (2015), was used in this study. A number of other self-report and performance based measures were administered, in order to examine convergent validity. For the current study, we only used those instruments that have previously been validated for use within the Italian context. These instruments are detailed below.

The original sample size of our Online Dataset was 287. We decided to exclude two participants: one because she was 17 years old, and one because she was non-Italian and resided in Italy less than 10 years (2.7 years). Our final sample included 285 participants from 18 to 68 years old (M = 26.4; SD = 7.0), 224 of which were women (78,6%). About 60% were university students (n = 166) and the other 40% was comprised of individuals with various occupations or unemployed.

Procedure

For Peer Review Only

Paper-and-pencil Dataset. These data were collected at two Italian universities, located in Milan and Rome. Prospective participants had been personally invited in class by the research assistant to volunteer about a study on psychology and interpersonal relationships. Inclusion/exclusion criteria were: (a) Italian citizenship, (b) fluent in the Italian language, and (c) not receiving psychiatric medications.

Online Dataset. This sample was recruited using flyers, social networking, and word of mouth; the volunteers were informed about a research study on the capability to recognize others' emotions. The data were collected using the "Google Form" service. Inclusion/exclusion criteria were the same as in the paper-and-pencil study.

Compliance with Ethical Standards

Experiment Participants. The current study used data retrieved from a previously published research study (paper-and-pencil dataset) and newly collected (online dataset) data. In both cases, the pertinent ethics committees of the universities involved in these projects (i.e., Sapienza University of Rome for the paper-and-pencil and University of Turin for the online study) gave their approval prior to beginning data collection.

Informed Consent. At the intake, prospective participants were told that participation was voluntary, that they could interrupt or end their participation at any time, and that questionnaires were anonymous. In line with the Helsinki declaration, all were asked to read and sign an informed consent statement prior to participating in the study. The participants were not offered any forms of incentive for participating, nor were they offered any monetary compensation for their participation.

Measures

For Peer Review Only

Both Datasets. All participants were administered the QCAE along with a number of other psychological scales. Ideally, to evaluate the convergent validity of a new empathy measure, one should try to use the most widely accepted measures of empathy, i.e., the IRI (Davis, 1980) and the EQ (Baron-Cohen, 1994). Although both had been validated in Italy (Albiero, Ingoglia, & Lo coco, 2006; Preti et al., 2011), 21 items of the QCAE were derived exactly from these two instruments (6 from the IRI and 15 from the EQ; Reniers, et al., 2011). For this reason, convergent validity was tested by focusing on constructs close to empathy, i.e., interpersonal competence, openness, extraversion, agreeableness, well-being, emotional regulation, and emotion recognition. Below (see Reliability and Validity Analyses) we describe more in detail the theoretical association of empathy to each of these constructs.

Questionnaire of Cognitive and Affective Empathy (QCAE; Reniers et al., 2011). The QCAE is an empathy measure composed of 31 items, rated on a 4-point Likert scale: strongly agree (1), slightly agree (2), slightly disagree (3), and strongly disagree (4). The translation of the QCAE was made in accordance with the classical translation-back-translation procedure (Geisinger, 2003): first, a bilingual individual translated the English original version into Italian language, then a second bilingual individual who were blind to the original QCAE version back-translated the italian version into English in order to identify potential discrepancies. The final, Italian QCAE version was eventually approved by two expert researchers who speak fluently both Italian and English.

Scores from each item were added to produce two subscales: Cognitive Empathy and Affective Empathy. These subscales are, in turn, composed of subcomponents. The subcomponents of Cognitive Empathy are *Perspective Taking* (PT) and *Online Simulation* (OS). PT measures the capability to put oneself in another person's shoes, while OS assesses attempt to

For Peer Review Only

put oneself in another person's position by imagining what that person is feeling and is likely to be used for future intentions. The subcomponents of Affective Empathy are *Emotion Contagion* (EC), *Proximal Responsivity* (PrR) and *Peripheral Responsivity* (PeR). EC is focused on the automatic mirroring of other's feelings. PrR is a measure of the emotional responsiveness to the feelings of others who are close within the social or affective subject's context. PeR is similar to PrR, however, its context is detached, such as experiencing empathy with protagonists in a film or a novel.

Paper-and-Pencil Dataset. In addition to the QCAE, participants included in this sample also completed the following questionnaires.

Interpersonal Competence Questionnaire (ICQ; Buhrmester et al., 1988). The ICQ is composed of 40 items measured on a 5-point Likert scale. The items make up five subscales: a) the ability to initiate relationships, b) the ability to assert displeasure with others, c) the ability to disclose personal information, d) the ability to provide emotional support and advice, and e) the ability to manage interpersonal conflict. Reliability and validity of ICQ scores was demonstrated by Buhrmester et al. (1988) and Giromini et al. (2015) for the original and the Italian versions respectively. In our sample, Cronbach alpha's were: .86 (Initiation Relationship), .77 (Emotional Support), .77 (Negative Assertion), .81 (Disclosure), and .78 (Conflict Management).[1]

NEO Five-Factor Inventory. (NEO-FFI; McCrae & Costa, 2004). The NEO-FFI is a short version of the Revised NEO Personality Inventory (NEO PI-R; Costa & McCrae, 1992), an instrument that measures personality traits of openness, conscientiousness, extraversion, agreeableness, and neuroticism. It is comprised of 60 items, all measured on a 5-point Likert scale. The Cronbach alpha's values in our sample were .68 (Openness), .72 (Conscientiousness), .63 (Extraversion), .62 (Agreeableness), and .75 (Neuroticism), which were similar to the ones reported

URL: http://mc.manuscriptcentral.com/JPersAssess Email: jpa_office@emich.edu

For Peer Review Only

by McCrae and Costa (2004) and those found in the Italian validation study of the NEO-FFI (Caprara et al., 2001).[1]

Psychological General Well-Being Index (PGWBI; Dupuy, 1977, 1984). The PGWBI is a 20-item self-report scale that assesses psychological well-being. Each item is measured on a 6-point Likert scale and the total score is broken down into six subscales: absence of anxiety, absence of depressed mood, positive well-being, self-control, general health and vitality. In this study we used the Italian version validated by Grossi et al. (2002). Cronbach alpha's were .84 (Absence of Anxiety), .76 (Absence of Depression), .82 (Positive well-being), .56 (Self-control), .59 (General health), .69 (Vitality), and .92 (Total PGWBI Score). Since the introduction of the original version of the PGWBI (Dupoy, 1984), many studies have used this instrument and provided support for its validity (e.g., Badia et al. 1996; Naughton & Wiklund, 1993).[1]

Difficulties in Emotion Regulation Scale (DERS; Gratz & Roemer, 2004). The DERS is a measure of difficulties in emotion regulation. It includes 36 items measured on a 5-point Likert scale, differentiating six areas of emotion regulation problems: a) non-acceptance of emotional responses, b) difficulties in engaging in goal-directed behavior, c) difficulties in controlling impulses, d) lack of emotional awareness, e) limited access to emotion regulation strategies, and f) lack of emotional clarity. Previous studies have demonstrated excellent psychometric properties for the DERS both in Italian (de Campora et al. 2014; Giovannini et al. 2014; Giromini et al., 2012, 2015) and foreign studies (e.g., Cho & Hong 2013; Dan-Glauser & Scherer, 2013; Gratz & Roemer, 2004; Mitsopoulou et al. 2013; Ruganci & Gencöz, 2010). In our sample, Cronbach alpha's were .86 (Nonacceptance), .86 (Goals), .87 (Impulse), .72 (Awareness), .90 (Strategies), .88 (Clarity), and .95 (Total DERS Score).[1]

For Peer Review Only

Online Dataset. In addition to the QCAE, participants in this study were also administered the following tests:

The Toronto Alexithymia Scale-20 items (TAS-20; Bagby, Taylor, Parker, & Loiselle, 1994). The TAS-20 is a self-report questionnaire composed of 20 items, rated on a 5-point Likert scale. In addition to the total TAS-20 score, three subscale scores are typically used, i.e., Difficulty Identifying Feelings (F1), Difficulty Describing Feelings (F2), and Externally Oriented Thinking (F3). In the original study by Bagby et al. (1994), the TAS-20 demonstrated acceptable internal consistency both for the total score (Cronbach's alpha = .81), and for each factor (Fl = .78, F2 = .75, F3 = .66). The Italian version of the TAS-20 (Bressi et al., 1996) also showed Cronbach alpha's values ranging from .52 to .77 in non-clinical sample. In our sample, Cronbach alpha's were .85 (Difficulty Identifying Feelings), .79 (Difficulty Describing Feelings), .67 (Externally Oriented Thinking), and .85 (Total score).

Reading the Mind in the Eyes-Test (RME-T; Baron-Cohen et al., 1997; Baron-Cohen, Wheelwright, Hill et al., 2001). The RME-T is used to assess emotion recognition; it includes 36 still pictures of the eye region, the person has to choose among four emotions that the pictures could represent; a link with the standardized glossary of the RME-T was present in the online administration. The aim of the test is to measure the frequency of matching a semantic definition to its expression in the picture and the score is calculated by the sum of correct responses. The Italian version of the RME-T was introduced by Vellante et al. (2013) who reported information on internal consistency, factor structure, and test-retest reliability of the Italian adaptation. The results of their study support the reliability of the Italian RME-T, although this instrument has produced low internal consistency indexes in other studies (Harkness et al., 2010; Khorashad et

For Peer Review Only

al., 2015; Olderbak et al., 2015; Ragsdale & Foley, 2011; Voracek & Dressler, 2006). In our study, Cronbach's alpha was relatively low (.32).

Statistical Analyses

Samples Homogeneity. Table 1 reports on the homogeneity of the samples composition. Both samples had a similar percentage of men and women, Phi = .05, $p = .18$. In both samples about three quarters were women. Conversely, a statistically significant difference emerged when examining the mean age of the two samples: Participants in the online dataset were significantly older than those in the paper-and-pencil sample, t (454.7) = 8.70, $p < .01$, $d = .67$ and mean ages were 22.6 vs. 26.4, respectively. We checked in the combined sample, whether age correlated with QCAE scores, which it did not: $|r| \leq .068$, $p \geq .074$. Furthermore, when we performed additional analyses (i.e., ANCOVA's) aimed at controlling for the impact of age on the mean differences between the paper-and-pencil and online formats, the results were virtually identical to those we obtained when the variable age was not controlled for. Likewise, because the paper-and-pencil sample only included university students while the online sample also included non-student participants, we performed additional analyses controlling for this possible confounding factor. After excluding all non-students from the combined dataset, we obtained similar results to those obtained when analyzing all available data. Thus, the comparison between the paper-and-pencil versus online administration was not notably affected by sample composition in terms of age or being a student or not.

Reliability and Validity Analyses. For both the paper-and-pencil and online datasets, we examined internal consistency and construct validity of QCAE scales. More in detail, QCAE scores' reliability was inspected via examination of Cronbach's alpha and item-scale correlations. Construct validity was tested by performing a confirmatory factor analysis (CFA) and by

For Peer Review Only

correlating QCAE scores to empathy-related constructs, such as interpersonal competence and personality traits such as agreeableness and openness (convergent validity). The comparison between paper-and-pencil vs. online QCAE scores was performed via t-test statistics, after testing CFA measurement invariance between the two formats. For both the correlational and t-test analyses, Holm-Bonferroni correction (Holm, 1979) was applied to correct for multiple testing.

As for the convergent validity, based on the previous literature on empathy and on our theoretical considerations, we expected that the Italian QCAE would correlate positively with psychological well-being, extraversion, openness, and agreeableness, but negatively with neuroticism and emotion dysregulation (Henry, Bailey, & Rendell, 2008). Indeed, extraverted individuals tend to be well-disposed and comfortable in human interactions (Costa & McCrae 1992). Conversely, neuroticism, anxiety, and depression probably decrease openness to social interactions, and the skills associated with facilitating them (Riemann & Allgöwer 1993).

Empathy was expected to positively correlate with interpersonal competence too, as empathic people are typically described as prone to get close to others in emotionally difficult times. Furthermore, because emotion regulation is considered to be one of the macro components involved in human empathy, especially in its development (Schipper & Petermann, 2013), we anticipated that empathy (especially affective empathy) would be negatively associated with difficulties in emotional regulation. Along the same lines, because existent literature suggests that alexithymia is correlated with the lack of empathy (Swart et al., 2009; Jonason & Krause, 2013), we also hypothesized that the QCAE (especially the cognitive empathy subscale) would negatively correlate with alexithymia.

Lastly, we also postulated that the QCAE would also correlate with the capability to recognize others' expressions and emotions, in that this skill is deemed to be a cognitive component too.

Results

Internal Consistency

Internal consistency of QCAE scores was estimated for both samples separately and for the combined sample (Table 2). Within the paper-and-pencil sample, internal consistency was adequate, with Cronbach alpha's ranging from .58 (*Proximal Responsivity* and *Peripheral Responsivity*) to .87 (*Perspective Taking*) for the subcomponents, and ≥ .77 for the *Cognitive Empathy and Affective Empathy* subscales and the *Total Score*. Similarly, within the online dataset, Cronbach alpha's ranged from .69 (*Peripheral Responsivity*) to .84 (*Perspective Taking*) for the subscales, and was ≥ .81 for the two subscales and total score. For *Proximal Responsivity* and *Peripheral Responsivity,* in the paper-and-pencil dataset Cronbach alpha's were .64 and .58, respectively; while in the online datasets, Cronbach alpha's were .69 for both the subscales.

Factor Structure

To factor analyze our QCAE data, we used Lisrel 8.72 (Jöreskog & Sörbom, 2005). Because these analyses aimed at testing whether the factor structure identified by Reniers et al. (2011) would also fit our Italian data, the same methodological approach utilized by Reniers et al. (2011) was used in this study, too. That is, we specified five latent variables (the five scales of the QCAE) and used the same item parcels utilized by the authors (for using item parceling rather than individual items in CFA, see Hall, Snell, & Foust, 1999; Little, Cunningham, Shahar, & Widaman, 2002; Nasser & Wisenbaker, 2003). Then, the same two models proposed – and tested via CFA – by Reniers et al. (2011) were tested. More specifically, in model 1 (M1), the five la-

For Peer Review Only

tent variables (i.e., the five QCAE subscales) were allowed to correlate with each other; in model 2 (M2), a hierarchical structure was tested, with Cognitive and Affective Empathy serving as second order factors (see Figure 1). Additionally – and differently from Reniers et al. (2011) – our study also tested a unidimensional model (UM) to provide us with a baseline referent model, to better evaluate M1 and M2.

The following goodness of fit statistics were taken under consideration, for all these three models. First, we looked at the χ^2, its associated p-value, and, most importantly, at the ratio between the χ^2 and its degrees of freedom (χ^2/df). According to Watkins (1989), a χ^2/df close to 2 reflects a good fit, and values lower than 5 indicate a quite promising fit. Next, we inspected the root mean square error of approximation (RMSEA) and its 90% confidence interval. Based on Browne and Cudeck (1993), we considered RMSEA values close to .05 to indicate a close fit, values close to .08 to indicate a fair fit, and values close to .10 to indicate a marginal fit. We then inspected the standardized root mean square residual (SRMR), whose values were expected to be close to or lower than .08 to indicate of a good fit (Hu & Bentler, 1980). Moreover, the comparative fit index (CFI) and non-normed fit index (NNFI) were inspected too, with their values being expected to be .90 or higher to indicate a good fit (Bentler & Bonett, 1980). Lastly, we also considered the Akaike's Information Criterion (AIC), whose values may be used to compare different models, as the lower the AIC, the better the fit of the model (Akaike, 1973).

The results of our CFAs are reported in Table 3. Based on the criteria described above, the UM did not provide an adequate fit, and therefore it was discharged. Conversely, both M1 and M2 fit relatively well our data. For example, both M1 and M2 produced RMSEA values below .10, SRMR values close to or lower than .08, and CFI and NNFI values above .90, in all samples under consideration. We thus compared M1 versus M2 by using the χ^2 test (i.e., by test-

ing the difference between the two χ^2 values) and by examining their AICs. The results of these additional analyses, presented in Table 4, indicate that M1 provided a significantly better fit than did M2, $\chi^2 \geq 19.9$, $p < .001$, and produced notably lower AIC values. Accordingly, it was concluded that M1 offered the best fit for our data (its factor loadings are reported in Table 5).

Convergent validity

In Table 6, convergent validity analyses are reported. Convergent validity with the ICQ, NEO-FFI, PGWBI, and DERS was calculated for the paper-and-pencil sample, while convergent validity with the TAS-20 and RME-T was calculated for the online sample. Below we discuss correlations that were statistically significant after Holm-Bonferroni correction (Holm, 1979).

As for the correlations of QCAE to ICQ, it is interesting to note that Cognitive Empathy and the Total QCAE scores correlated positively with all ICQ scales, $r \geq .23$. Moreover, Affective Empathy correlated with the ICQ *Emotional Support* scale only, $r = .27$. Of all QCAE scales, *Emotion Contagion* was the only one that did not correlate with any of the ICQ scales.

Similarly, the correlations between the QCAE and NEO-FFI revealed a different pattern for Cognitive versus Affective Empathy. Cognitive Empathy correlated positively with *Extraversion* ($r = .26$), *Openness* ($r = .30$), and *Conscientiousness* ($r = .33$), and negatively with *Neuroticism* ($r = -.27$). Furthermore, Affective Empathy correlated only with *Neuroticism* ($r = .32$). Noteworthy, the effect size of the relationship between Emotion Contagion and Neuroticism was medium to large, i.e., $r = -.41$. All other correlations between the QCAE and NEO FFI consisted, at maximum, of a medium effect sizes.

The total well-being score (PGWBI Total) did not correlate with Total QCAE score, but correlated positively with Cognitive Empathy ($r = .17$), and negatively with Affective Empathy ($r = -.17$). Because of the two correlations are exactly the opposite, they cancel each other out in

For Peer Review Only

the final correlation between the Total QCAE score and the PGWBI Total. Again, when looking at the QCAE subcomponents, *Emotion Contagion* produced the strongest correlation with the total well-being score, r = -.23.

The Total QCAE score produced significant correlations with the total DERS (r = -.19) and total TAS-20 (r = -.27) scores. However, while Cognitive Empathy correlated r = -.34 with the total DERS score and r = -.35 with the total TAS-20 score, Affective Empathy did not correlate with these two. *Emotion Contagion* was the only one that produced positive correlations with DERS and TAS-20 scales and/or subscales.

Finally, the QCAE did not produce statistically significant correlations with the RME-T.

Comparison between QCAE scores from Paper-and-Pencil and Online administrations

Prior to comparing QCAE scores from the paper-and-pencil versus the online administration, we tested CFA measurement invariance between the two formats. Because M1 provided the best fit our data, M1 only was analyzed for structural invariance. These analyses were performed across four steps, in line with previous research in the field assessment (e.g., Beaujean, Freeman, Youngstrom & Carlson, 2012). That is, first, configural invariance assessed if the factor model was invariant across the two groups/formats. Next, metric invariance investigated if the factor loadings for QCAE parcels (Table 5) were the same in both groups/formats. Third, scalar invariance was tested by constraining all the scales' origins (i.e., intercepts) across the two groups/formats. Finally, invariant unique variance analyzed the invariance of the unique residual variances across the two groups/formats. The results of these analyses, reported in Table 7, suggest that the paper-and-pencil and online formats were structurally invariant. Indeed, all models were adequate, and no notable differences from one step to another were observed.

URL: http://mc.manuscriptcentral.com/JPersAssess Email: jpa_office@emich.edu

Accordingly, we next performed a series of t-tests to compare the scores produced by the two formats. As shown in Table 8, the online version produced statistically significantly higher QCAE scores than the paper-and-pencil version, with a *small* or *small to medium* effect size (Cohen, 1988). This difference was more evident with the female than with the male samples. Also noteworthy, although Reniers et al. (2011) did not report detailed, descriptive statistics concerning their samples' QCAE scores, they did report the average scores (and relative standard errors) of men and women on the Cognitive Empathy and Affective Empathy QCAE scales. More specifically, men had a mean of 56.1 (*SE* = .5) on Cognitive Empathy and a mean of 32.3 (*SE* = .3) on Affective Empathy, and women had a mean of 59.4 (*SE* = .3) on Cognitive Empathy and a mean of 36.8 (*SE* = .2) on Affective Empathy. These values – which were obtained by Reniers et al. (2011) via online administration – are markedly similar, nearly identical to those observed in this study, when considering the online sample data.

Discussion

Over the last 20 years, the empathy construct has been refined, notably the distinction between cognitive and affective components of empathy. In line with these refinements, Reniers et al. (2011) developed the Questionnaire of Cognitive and Affective Empathy (QCAE), a 31-item self-report measure of cognitive and affective empathy. The main purpose of our study was to examine internal consistency, factor structure, and convergent validity of an Italian version of the QCAE. Furthermore, we also compared the average scores obtained by a paper-and-pencil version with online administration of the QCAE.

In terms of internal consistency, all QCAE scales from both the paper-and-pencil and online versions produced Cronbach's alpha values above .70, except for *Proximal Responsivity* and *Peripheral Responsivity*, which produced Cronbach's alpha scores between .58 and .69.

For Peer Review Only

Thus, both the paper-and-pencil and online formats produced similar internal consistency results, and these results are comparable to those reported by Reniers et al. (2011). It is noteworthy that in Reniers et al.'s (2011) study *Proximal Responsivity* and *Peripheral Responsivity* also demonstrated the lowest Cronbach's alpha values of all QCAE subscales (.70 and .65, respectively). Although many statisticians criticize the idea that Cronbach's alpha values below .70 reflect lack of reliability for the scales under investigation (e.g., John & Soto, 2007; Sijtsma, 2009), future studies should pay particular attention to the reliability of scores from these two subscales. Perhaps, a tentative explanation for these relatively low reliability indices may be that these two subcomponents, along with *Emotion Contagion*, are the ones with the lowest number of items (i.e., 4 each).

The results of our confirmatory factor analysis suggest that both the data from the paper-and-pencil and those from the online dataset fit the models proposed by Reniers et al. (2011) relatively well, and certainly better than did the unidimensional model. Our study thus suggests that the QCAE shows structural validity and factorial stability across different Western cultures and languages, regardless of administration format. On the other hand, since the QCAE has only been studied in Western populations, additional validation research in non-Western samples is necessary. Furthermore, and perhaps more importantly, because our goal was to test whether the model(s) proposed by Reniers et al. (2011) would also fit our Italian data, we decided to use the same methodological approach that they used in their study. That is, we decided to perform our CFAs on item parcels rather than on individual items. On one hand, this approach ensured that any potential discrepancies in the CFA results of our vs. Reniers et al.'s (2011) studies could not be due to the analytic strategies being different from one study to another. On the other hand, however, because the QCAE items are scored on a 4-point Likert scale, future studies performing

For Peer Review Only

CFAs on all items (e.g., by using an appropriate estimator with robust standard errors) would probably be beneficial.

Our convergent validity analyses revealed some interesting patterns. Cognitive Empathy correlated positively with interpersonal competence (ICQ), amiable personality traits of the NEO-FFI (i.e., extraversion, openness, and conscientiousness), and psychological well-being (PGWBI), and negatively with alexithymia (TAS-20), neuroticism (NEO-FFI), and difficulties with emotion regulation (DERS). Those correlations were anticipated, as they are largely in line with recent findings that associate dysfunction in social functioning with risk of a variety of psychopathological conditions, such as psychosis (Henry, Bailey, & Rendell, 2008).

Conversely, Affective Empathy produced a more complex pattern of correlations, which cannot be understood without looking at the correlations produced by its subscales: While *Proximal Responsivity* correlated positively with interpersonal competence and amiable personality traits (i.e., extraversion, openness, and agreeableness), *Emotion Contagion* – and to a lesser extent *Peripheral Responsivity* – correlated negatively with well-being (PGWBI), and positively with alexithymia (TAS-20), neuroticism (NEO-FFI), and difficulties in emotion regulation (DERS). Given the negative correlations between QCAE Affective Empathy and well-being, considering affective empathy as a resource does not appear to be a foregone conclusion. In fact, high levels of emotion contagion had recently been associated with some pathological conditions. Weisbuch et al. (2011) found an increased risk for eating disorders in young women susceptible to emotion contagion. Also high levels of emotion contagion were found in patients with schizophrenia (Horan et al., 2015) and with difficulties in emotional regulation (Miguel et al., 2016). Combined with previous literature, our findings thus suggest that emotion contagion could be associated with psychological vulnerability.

For Peer Review Only

Taken together, our convergent validity results also support the cross-cultural applicability of the QCAE. Indeed, the QCAE total score produced positive and statistically significant correlations with instruments measuring constructs related to empathy (i.e., interpersonal competence and amiable personality traits) and negative correlations with difficulties in recognition and regulation of emotions (i.e., alexithymia and emotional dysregulation). Conversely, the correlation of the QCAE to emotion recognition was nonsignificant. Given that emotion recognition and empathy are only partially overlapping constructs, the relatively weak correlation between QCAE and RME-T is not unexpected, but rather suggests that emotion recognition is probably necessary, but not sufficient to empathize with others. Furthermore, it should be noted that while the QCAE is a self-report measure, the RME-T is rather a performance-based instrument. As such, it is not too surprising that the two instruments do not correlate strongly with each other (Mihura et al., 2013).

One of the most interesting results of our study, in our opinion, is that when compared to the standard, paper-and-pencil format, the online administration format produced significantly higher QCAE scores. Based on our post-hoc analyses controlling for age and student status (i.e., being a student vs. not being a student), it is unlikely that these differences may be accounted for simply by demographic heterogeneity across the two samples. Perhaps, a better explanation for these findings may be ascribed to self-selection bias (e.g., participants had not been personally invited by the research assistant to volunteer for the study) and under-coverage in online surveys (Bethlehem, 2010). Said differently, it is possible that those who decided by themselves to volunteer in the online group were more interested in knowing about emotions and empathy compared to the paper-and-pencil sample (who was explicitly asked to volunteer by an assistant) – a

URL: http://mc.manuscriptcentral.com/JPersAssess Email: jpa_office@emich.edu

For Peer Review Only

characteristic, that is typical of empathetic individuals. Future QCAE research might further inspect whether online administrations produce higher scores than paper-and-pencil format.

Although our findings provide initial support for the cross-cultural applicability of the QCAE, some of our study's limitations deserve mentioning. First, our two samples are far from being representative of the general Italian population, and some demographic and sample size differences between the paper-and-pencil and online samples make it difficult to rule out that the two samples scored differently on the QCAE for some uncontrolled reasons. For these reasons, our findings still need to be further replicated with other samples too. Second, but somehow related to this first point, future studies should attempt to control for many other variables that we could not control for in our study, such as socioeconomic status, marital status, etc. Third, one of the instruments we used to test convergent validity, the RME-T, had a very low internal reliability. As such, the generalizability of its results to other studies is difficult to evaluate. Fourth, we did not examine divergent validity or test-retest stability, which are important to better estimate the validity and reliability of our QCAE scores.

Despite these limitations, our study is the first to investigate the reliability and validity of the QCAE in Italy, and to compare QCAE scores obtained with paper-and-pencil versus online administration formats. We found evidence that the Italian version of the QCAE has sound psychometric properties. We showed the QCAE had adequate internal reliability, factorial stability and convergent validity. As such, the instrument holds promise as an easy to administer self-report tool for the assessment of the cognitive and affective components of empathy.

Footnotes

[1] These Cronbach alpha's values were previously reported by Giromini et al. (2015).

For Peer Review Only

References

American Psychiatric Association, & American Psychiatric Association. DSM-5 Task Force (2013). *Diagnostic and statistical manual of mental disorders: DSM-5™*. *Arlington, VA: American Psychiatric Publishing*.

Badia, X., Gutiérrez, F., Wiklund, I., & Alonso, J. (1996). Validity and reliability of the Spanish version of the psychological general wellbeing index. *Quality of Life Research, 5*(1), 101–108.

Bagby, R. M., Parker, J. D., & Taylor, G. J. (1994). The twenty-item Toronto Alexithymia Scale—I. Item selection and cross-validation of the factor structure. *Journal of psychosomatic research*, *38*(1), 23-32.

Baron-Cohen, S. (2005). The empathizing system: A revision of the 1994 model of the Min-dreading System. In: Ellis, B. and Bjorklund, D. (Eds.), Origins of the Social Mind. Guilford, New York, pp. 468–492.

Baron-Cohen, S., & Wheelwright, S. (2004). The empathy quotient: an investigation of adults with Asperger syndrome or high functioning autism, and normal sex differences. *Journal of autism and developmental disorders*, *34*(2), 163-175.

Bateson, C. D. (2009). *These things called empathy: Eight related but distinct phenomena*. In J. Decety & W. Ickes (Eds), The social neuroscience of empathy (pp. 3-15). Cambridge (Mass.): MIT Press.

Bentler, P. M., & Bonett, D. G. (1980). Significance tests and goodness of fit in the analysis of covariance structures. *Psychological bulletin*, *88*(3), 588.

Bethlehem, J. (2010). Selection bias in web surveys. *International Statistical Review*, *78*(2), 161-188.

For Peer Review Only

Blair, R. J. R. (2005). Responding to the emotions of others: Dissociating forms of empathy through the study of typical and psychiatric populations. *Consciousness and Cognition: An International Journal, 14*, 698–718.

Bohart, A. C., & Greenberg, L. S. (1997). Empathy and psychotherapy: An introductory overview. In A. C. Bohart & L. S. Greenberg (Eds.), *Empathy reconsidered: New directions in psychotherapy* (pp. 1–39). Washington, DC: American Psychological Association.

Bressi, C., Taylor, G., Parker, J., Bressi, S., Brambilla, V., Aguglia, E., ... & Todarello, O. (1996). Cross validation of the factor structure of the 20-item Toronto Alexithymia Scale: an Italian multicenter study. *Journal of psychosomatic research*, *41*(6), 551-559.

Brown, M. M., Thibodeau, R. B., Pierucci, J. M., & Gilpin, A. T. (2017). Supporting the development of empathy: The role of theory of mind and fantasy orientation. *Social Development,* (November), 1–14.

Browne, M. W., & Cudeck, R. (1993). Alternative ways of assessing model fit. *Sage focus editions*, *154*, 136-136.

Buhrmester, D., Furman, W., Wittenberg, M. T., & Reis, H. T. (1988). Five domains of interpersonal competence in peer relationships. *Journal of Personality and Social Psychology, 55*(6), 991–1008. doi:10.1037/0022-3514.55.6.991.

Caprara, G. V., Barbaranelli, C., Hahn, R., & Comrey, A. L. (2001). Factor analyses of the NEO-PI-R inventory and the Comrey personality scales in Italy and the United States. *Personality and Individual Differences, 30,* 217–228.

Cho, Y., & Hong, S. (2013). The new factor structure of the Korean version of the Difficulties in Emotion Regulation Scale (K-DERS) incorporating method factor. *Measurement and Evaluation in Counseling and Development*, *46*(3), 192-201.

For Peer Review Only

Cohen, J. (1988). *Statistical Power Analysis for the Behavioral Sciences,* 2nd ed. Hillsdale, New Jersey: Erlbaum.

Costa, P. T., & McCrae, R. R. (1992). Revised NEO personality inventory (NEO-PI-R) and NEO five-factor inventory (NEO-FFI) professional manual. Odessa: Psychological Assessment Resources.

Damasio, A.R. (1994). *Descartes' Error: Emotion, Reason, and the Human Brain.* Putnam, New York.

Dan-Glauser, E. S., & Scherer, K. R. (2013). The Difficulties in Emotion Regulation Scale (DERS): Factor structure and consistency of a French translation. *Swiss Journal of Psychology*, *72*(1), 5.

Davis, M. H. (1980). A multidimensional approach to individual differences in empathy. *JSAS Catalog of Selected Documents in Psychology 10*, 85.

de Campora, G., Giromini, L., Larciprete, G., Volsi, V. L., & Zavattini, G. C. (2014). The impact of maternal overweight and emotion regulation on early eating behaviors. *Eating behaviors*, *15*(3), 403-409.

Dupuy, H. J. (1977). The general well-being schedule. In I. McDowell & C. Newell (Eds.), *Measuring health: A guide to rating scales and questionnaire* (2nd ed., pp. 206–213). USA: Oxford University Press.

Dupuy, H. J. (1984). The psychological general well-being (PGWB) index. In N. Wenger (Ed.), *Assessment of quality of life in clinical trials of cardiovascular therapies* (pp. 170–183). New York: Le Jacq.

For Peer Review Only

Feather, N. T., & Nairn, K. (2005). Resentment, envy, schadenfreude, and sympathy: Effects of own and other's deserved or undeserved status. *Australian Journal of Psychology*, *57*(2), 87-102.

Force, P. T. (2006). Psychodynamic diagnostic manual.

Frith, U., & Frith, C. D. (2003). Development and neurophysiology of mentalizing. *Philosophical Transactions of the Royal Society B: Biological Sciences*, *358*(1431), 459-473.

Geisinger, K. F. (2003). Testing and assessment in cross-cultural psychology. In J. R. Graham, J. A. Naglieri, & I. B. Weiner (Eds.), *Handbook of psychology* (Vol. 10). Assessment psychology (pp.95–118). New Jersey: Wiley.

Geisinger, K. F. (2003). Testing and assessment in cross-cultural psychology. In J. R. Graham, J. A. Naglieri, & I. B. Weiner (Eds.), *Handbook of psychology* (Vol. 10). Assessment psychology (pp.95–118). New Jersey: Wiley.

Gibbons, S. B. (2011). Understanding empathy as a complex construct: A review of the literature. *Clinical Social Work Journal*, *39*(3), 243-252.

Giovannini, C., Giromini, L., Bonalume, L., Tagini, A., Lang, M., & Amadei, G. (2014). The Italian five facetmindfulness questionnaire: a contribution to its validity and reliability. *Journal of Psychopathology and Behavioral Assessment, 36,* 415–423. doi:10.1007/s10862-013-9403-0.

Giromini, L., Brusadelli, E., Di Noto, B., Grasso, R., & Lang, M. (2015). Measuring psychological mindedness: validity, reliability, and relationship with psychopathology of an Italian version of the balanced index of psychological mindedness. *Psychoanalytic Psychotherapy, 29*(1), 70–87. doi:10.1080/02668734.2015.1006666.

For Peer Review Only

Giromini, L., Velotti, P., De Campora, G., Bonalume, L., & Zavattini, G. C. (2012). Cultural adaptation of the difficulties in emotion regulation scale: reliability and validity of an Italian version. *Journal of Clinical Psychology, 68*(9), 989–1007. doi:10.1002/jclp.21876.

Gratz, K. L., & Roemer, L. (2004). Multidimensional assessment of emotion regulation and dysregulation: Development, factor structure, and initial validation of the difficulties in emotion regulation scale. *Journal of Psychopathology and Behavioral Assessment*, *26*(1), 41-54.

Grossi, E., Masconi, P., Groth, N., Nievo, M., & Apolone, G. (2002). Il questionario psychological general well being. Versione Italiana. Versione 0.2. Milan: Marzo.

Hall, R. J., Snell, A. F., & Foust, M. S. (1999). Item parceling strategies in SEM: Investigating the subtle effects of unmodeled secondary constructs. *Organizational Research Methods*, *2*(3), 233-256.

Harkness, K. L., Jacobson, J. A., Duong, D., & Sabbagh, M. A. (2010). Mental state decoding in past major depression: Effect of sad versus happy mood induction. *Cognition and Emotion*, *24*(3), 497-513.

Henry, J. D., Bailey, P. E., & Rendell, P. G. (2008). Empathy, social functioning and schizotypy. *Psychiatry research, 160*(1), 15-22.

Hoffman, M.L. (1984) Interaction of affect and cognition in empathy. In: Izard, C., Kagan, J. and Zajonc, R. (Eds.), *Emotions, Cognition, and Behavior.* Cambridge University Press, New York, pp. 103–131.

Holm, S. (1979). A simple sequentially rejective multiple test procedure. *Scandinavian journal of statistics*, 65-70.

For Peer Review Only

Horan, W. P., Reise, S. P., Kern, R. S., Lee, J., Penn, D. L., & Green, M. F. (2015). Structure and correlates of self-reported empathy in schizophrenia. *Journal of psychiatric research*, *66*, 60-66.

Hu, L. T., & Bentler, P. M. (1999). Cutoff criteria for fit indexes in covariance structure analysis: Conventional criteria versus new alternatives. *Structural equation modeling: a multidisciplinary journal*, *6*(1), 1-55.

John, O. P., & Soto, C. J. (2007). The importance of being valid. *Handbook of research methods in personality psychology*, 461-494.

Jolliffe, D., & Farrington, D. P. (2004). Empathy and offending: A systematic review and meta-analysis. *Aggression and violent behavior*, *9*(5), 441-476.

Jonason, P. K., & Krause, L. (2013). The emotional deficits associated with the Dark Triad traits: Cognitive empathy, affective empathy, and alexithymia. *Personality and Individual Differences, 55*(5), 532-537.

Jöreskog, K. G., & Sörbom, D. (2005). *LISREL 8.72: Interactive LISREL for MS windows*. Lincolnwood: Scientific Software International, Inc.

Kanske, P., Böckler, A., Trautwein, F. M., & Singer, T. (2015). Dissecting the social brain: Introducing the EmpaToM to reveal distinct neural networks and brain-behavior relations for empathy and Theory of Mind. *NeuroImage, 122,* 6–19. https://doi.org/10.1016/j.neuroimage.2015.07.082

Khorashad, B. S., Baron-Cohen, S., Roshan, G. M., Kazemian, M., Khazai, L., Aghili, Z., ... & Afkhamizadeh, M. (2015). The "Reading the Mind in the Eyes" test: investigation of psychometric properties and test–retest reliability of the persian version. *Journal of autism and developmental disorders*, *45*(9), 2651-2666.

For Peer Review Only

Lawrence, E. J., Shaw, P., Baker, D., Baron-Cohen, S., & David, A. S. (2004). Measuring empathy: reliability and validity of the Empathy Quotient. *Psychological medicine*, *34*(05), 911-920.

Leiberg, S., & Anders, S. (2006). The multiple facets of empathy: a survey of theory and evidence. *Progress in brain research, 156,* 419-440.

Lipps, T. (1903). Einfühlung, innere Nachahmung und Organenempfindungen.

Little, T. D., Cunningham, W. A., Shahar, G., & Widaman, K. F. (2002). To parcel or not to parcel: Exploring the question, weighing the merits. *Structural equation modeling*, *9*(2), 151-173.

Swart, M., Kortekaas, R., & Aleman, A. (2009). Dealing with feelings: characterization of trait alexithymia on emotion regulation strategies and cognitive-emotional processing. *PLoS One*, *4*(6), e5751.

McCrae, R. R., & Costa, P. T. (2004). A contemplated revision of the NEO five-factor inventory. *Personality and Individual Differences, 36*(3), 587–596. doi:10.1016/S0191-8869(03)00118-1.

Miguel, F. K., Giromini, L., Colombarolli, M. S., Zuanazzi, A. C., & Zennaro, A. (2016). A Brazilian Investigation of the 36-and 16-Item Difficulties in Emotion Regulation Scales. *Journal of Clinical Psychology, advance online publication*.

Mihura, J. L., Meyer, G. J., Dumitrascu, N., & Bombel, G. (2013). The validity of individual Rorschach variables: Systematic reviews and meta-analyses of the comprehensive system. *Psychological Bulletin*, 139(3) 548 – 605.

Mitsopoulou, E., Kafetsios, K., Karademas, E., Papastefanakis, E., & Simos, P. G. (2013). The Greek version of the difficulties in emotion regulation scale: testing the factor structure, reliability and validity in an adult community sample. *Journal of Psychopathology and Behavioral Assessment*, *35*(1), 123-131.

For Peer Review Only

Nasser, F., & Wisenbaker, J. (2003). A Monte Carlo study investigating the impact of item parceling on measures of fit in confirmatory factor analysis. *Educational and Psychological Measurement*, *63*(5), 729-757.

Naughton, M. J., & Wiklund, I. (1993). A critical review of dimension-specific measures of health-related quality of life in cross-cultural research. *Quality of life research*, *2*(6), 397-432.

Olderbak, S., Wilhelm, O., Olaru, G., Geiger, M., Brenneman, M. W., & Roberts, R. D. (2015). A psychometric analysis of the reading the mind in the eyes test: toward a brief form for research and applied settings. *Frontiers in psychology*, *6*, 1503.

Preston, S.D., & de Waal, F.B. (2002). Empathy: its ultimate and proximate bases. *Behav. Brain Sci., 25*, 1–20.

Preti, A., Vellante, M., Baron-Cohen, S., Zucca, G., Petretto, D. R., & Masala, C. (2011). The Empathy Quotient: A cross-cultural comparison of the Italian version. *Cognitive Neuropsychiatry, 16*(1), 50-70.

Ragsdale, G., & Foley, R. A. (2011). A maternal influence on Reading the mind in the Eyes mediated by executive function: differential parental influences on full and half-siblings. *PloS one*, *6*(8), e23236.

Reniers, R. L., Corcoran, R., Drake, R., Shryane, N. M., & Völlm, B. A. (2011). The QCAE: A questionnaire of cognitive and affective empathy. *Journal of personality assessment*, *93*(1), 84-95.

Ritter, P., Lorig, K., Laurent, D., & Matthews, K. (2004). Internet versus mailed questionnaires: A randomized comparison. *Journal of Medical Internet Research, 6(3)*, e29.

For Peer Review Only

Riva, G., Teruzzi, T., & Anolli, L. (2003). The use of the Internet in psychological research: Comparison of online and offline questionnaires. *Cyberpsychology & Behavior: The Impact of the Internet, Multimedia and Virtual Reality on Behavior and Society, 6(1)*, 73–80.

Rizzolatti, G., Fadiga, L., Gallese, V., & Fogassi, L. (1996) Premotor cortex and the recognition of motor actions. *Brain Res. Cogn. Brain Res., 3,* 131–141.

Rogers, C. (1959) A theory of therapy, personality, and interpersonal relationships as developed in the client-centered framework. In: Koch, J.S. (Ed.) *Psychology: a Study of a Science*: Vol. 3. Formulations of the Person in the Social Context. McGraw-Hill, New York, pp. 184–256.

Rugancı, R. N., & Gençöz, T. (2010). Psychometric properties of a Turkish version of the Difficulties in Emotion Regulation Scale. *Journal of Clinical Psychology*, *66*(4), 442-455.

Schipper, M., & Petermann, F. (2013). Relating empathy and emotion regulation: Do deficits in empathy trigger emotion dysregulation? *Social neuroscience, 8*(1), 101-107.

Sijtsma, K. (2009). On the use, the misuse, and the very limited usefulness of Cronbach's alpha. *Psychometrika*, *74*(1), 107.

Skitka, L. J., & Sargis, E. G. (2006). The Internet as psychological laboratory. *Annu. Rev. Psychol.*, *57*, 529-555.

Trivers, R. L. (1971). The evolution of reciprocal altruism. *The Quarterly review of biology, 46*(1), 35-57.

Vellante, M., Baron-Cohen, S., Melis, M., Marrone, M., Petretto, D. R., Masala, C., & Preti, A. (2013). The "Reading the Mind in the Eyes" test: systematic review of psychometric properties and a validation study in Italy. *Cognitive neuropsychiatry*, *18*(4), 326-354.

For Peer Review Only

Völlm, B. A., Taylor, A. N., Richardson, P., Corcoran, R., Stirling, J., McKie, S., ... & Elliott, R. (2006). Neuronal correlates of theory of mind and empathy: a functional magnetic resonance imaging study in a nonverbal task. *Neuroimage*, *29*(1), 90-98.

Voracek, M., & Dressler, S. G. (2006). Lack of correlation between digit ratio (2D: 4D) and Baron-Cohen's "Reading the Mind in the Eyes" test, empathy, systemising, and autism-spectrum quotients in a general population sample. *Personality and Individual Differences*, *41*(8), 1481-1491.

Weisbuch, M., Ambady, N., Slepian, M. L., & Jimerson, D. C. (2011). Emotion contagion moderates the relationship between emotionally-negative families and abnormal eating behavior. *International Journal of Eating Disorders*, *44*(8), 716-720.

For Peer Review Only

Table 1. Composition of the Samples.

	Paper-and-pencil Dataset	Online Dataset	Combined Dataset
Gender, $\phi = .05$, $p = .18$			
Male	105 (25.9%)	61 (21.4%)	166 (24.0%)
Female	301 (74.1%)	224 (78.6%)	525 (76.0%)
Age, $t(454.7) = 8.70$, $p < .01$, $d = .67$*			
M	22.55	26.41	24.16
SD	4.61	7.01	6.03

* Because homoscedasticity could not be assumed, the Welch–Satterthwaite method was used to adjust degrees of freedom.

For Peer Review Only

ONLINE VS. PAPER-AND-PENCIL ITALIAN QCAE

Table 2. Internal Consistency Analyses.

QCAE Scale	No. of items	Paper-and-pencil Dataset		Online Dataset		Combined Dataset	
		Cronbach's alpha	Range of item-total correlations	Cronbach's alpha	Range of item-total correlations	Cronbach's alpha	Range of item-total correlations
Perspective Taking	10	.87	.62 – .71	.84	.54 – .72	.86	.60 – .71
Online Simulation	9	.83	.52 – .75	.78	.50 – .69	.82	.51 – .72
Emotion Contagion	4	.73	.69 – .79	.76	.69 – .81	.74	.69 – .80
Proximal Responsivity	4	.64	.57 – .77	.69	.60 – .83	.67	.59 – .80
Peripheral Responsivity	4	.58	.45 – .80	.69	.48 – .85	.63	.47 – .82
Cognitive Empathy	19	.89	.44 – .68	.84	.33 – .65	.87	.40 – .64
Affective Empathy	12	.77	.26 – .66	.81	.31 – .69	.79	.29 – .67
Total Score	31	.87	.26 – .60	.86	.31 – .61	.87	.29 – .60

For Peer Review Only

ONLINE VS. PAPER-AND-PENCIL ITALIAN QCAE

Table 3. Goodness of Fit Indices for a Univariate Model, and for Models 1 and 2.

	Paper-and-pencil Dataset			Online Dataset			Combined Dataset		
	UM	M 1	M 2	UM	M 1	M 2	UM	M 1	M 2
χ^2	1400.55	257.67	292.85	1192.13	268.99	288.89	2489.13	419.85	477.78
df	90	80	85	90	80	85	90	80	85
χ^2 p-value	< 0.01	< 0.01	< 0.01	< 0.01	< 0.01	< 0.01	< 0.01	< 0.01	< 0.01
χ^2/df	15.56	3.22	3.45	13.26	3.36	3.40	27.66	5.25	5.62
RMSEA	.19	.07	.08	.21	.09	.09	.20	.08	.08
RMSEA 90% CI	.18–.20	.06–.08	.07–.09	.20–.22	.08–.10	.08–.10	.19–.20	.07–.09	.08–.09
SRMR	.13	.07	.08	.15	.08	.09	.13	.07	.08
CFI	.78	.96	.95	.70	.93	.92	.76	.95	.95
NNFI	.75	.94	.94	.65	.91	.91	.72	.94	.94
AIC	1460.55	337.67	362.26	1252.13	348.99	358.89	2549.13	499.85	547.78

UM = Univariate model; M1 = Model 1; M2 = Model 2; RMSEA = root mean square error of approximation; SRMR = standardized root mean square residual; CFI = compared fit index; NNFI = non-normed fit index; AIC = Akaike's Information Criterion.

For Peer Review Only

Table 4. Comparison between Model 1 and Model 2.

	Paper-and-pencil Dataset	Online Dataset	Combined Dataset
Δ(M2 – M1) χ^2	35.18	19.9	57.93
Δ(M2 – M1) df	5	5	5
p-value Δ(M2 – M1) χ^2	< 0.001	< 0.001	< 0.001
Δ(M2 – M1) AIC	24.59	9.90	47.93

M1 = Model 1; M2 = Model 2

For Peer Review Only

Table 5. Factor Loadings for Reniers et al.'s (2011) QCAE Parcels, as Obtained from our CFA – M1.

Variable	Paper-and-pencil Dataset	Online Dataset	Combined Dataset
Perspective Taking			
P 11	.73	.80	.76
P 12	.77	.75	.77
P 13	.77	.69	.74
P 14	.78	.69	.76
P 15	.75	.74	.75
Online Simulation			
P 21	.72	.72	.72
P 22	.80	.80	.79
P 23	.77	.72	.76
P 24	.76	.62	.71
Emotion Contagion			
P 31	.71	.68	.69
P 32	.71	.78	.75
Proximal Responsivity			
P 41	.66	.68	.68
P 42	.69	.71	.70
Peripheral Responsivity			
P 51	.59	.71	.62
P 52	.72	.83	.79

Note. The labels "P 11, P 12, …, P 52" refer to the same item parcels utilized by Reniers et al. (2011)

For Peer Review Only

ONLINE VS. PAPER-AND-PENCIL ITALIAN QCAE

Table 6. Convergent Validity Analyses.

	Perspective Taking	Online Simulation	Emotion Contagion	Proximal Responsivity	Peripheral Responsivity	Cognitive Empathy	Affective Empathy	Total Score
Paper-and-pencil Dataset								
ICQ (n = 407)								
IR	.35**	.15	-.10	.17	.03	.30**	.04	.24**
ES	.52**	.45**	.01	.45**	.19*	.57**	.27**	.56**
NA	.41**	.17	-.12	.15	.01	.35**	.01	.26**
DC	.26**	.13	.03	.18*	.12	.23**	.15	.24**
CM	.26**	.43**	.03	.25**	.03	.40**	.13	.36**
NEO-FFI (n = 407)								
Neuroticism	-.25**	-.22**	.41**	.10	.21**	-.27**	.32**	-.05
Extraversion	.29**	.14	-.06	.19*	.07	.26**	.08	.23**
Openness	.31**	.19*	-.05	.25**	.22**	.30**	.17	.31**
Agreeableness	-.02	.31**	.07	.24**	.11	.16	.18*	.21**
Conscientiousness	.29**	.29**	.01	.15	.09	.33**	.11	.30**
PGWBI (n = 407)								
AA	.08	.11	-.23**	-.04	-.15	.11	-.19*	-.01
AD	.16	.16	-.14	.04	-.07	.19*	-.08	.10
PWB	.11	.09	-.19*	.01	-.07	.12	-.11	.03
SC	.17	.18*	-.19*	.02	-.07	.20**	-.11	.10
GH	.16	.13	-.11	.03	-.08	.17	-.07	.09
VIT	.08	.09	-.21**	-.07	-.14	.10	-.18*	-.02
Total	.15	.15	-.23**	-.01	-.13	.17	-.17	.05
DERS (n = 407)								
Nonacceptance	-.16	-.16	.24**	.01	.04	-.19*	.13	-.08
Goals	-.04	-.09	.25**	.10	.08	-.07	.19*	.04
Impulse	-.23**	-.32**	.23**	-.05	.06	-.32**	.11	-.19*

For Peer Review Only

Awareness	-.34**	-.26**	-.06	-.29**	-.19*	-.35**	-.23**	-.37**
Strategies	-.21**	-.25**	.29**	-.02	.08	-.26**	.16	-.12
Clarity	-.30**	-.22**	.20**	-.04	.01	-.30**	.08	-.19*
Total	-.28**	-.30**	.28**	-.05	.03	-.34**	.12	-.19*
Online Dataset								
TAS-20 (n = 285)								
DIF	-.16	-.15	.29**	.03	.00	-.18	.14	-.04
DDF	-.27**	-.13	.06	-.17	-.14	-.25**	-.10	-.22*
EOT	-.33**	-.38**	-.05	-.31**	-.27**	-.43**	-.26**	-.43**
Total	-.31**	-.27**	.15	-.17	-.16	-.35**	-.06	-.27**
RME-T (n = 282)								
Total	.16	.13	.00	.08	.08	.18	.07	.16

ICQ = Interpersonal Competence Questionnaire; IR = Initiation Relationship; ES = Emotional Support; NA = Negative Assertion; DC = Disclosure; CM = Conflict Management; NEOFFI = NEO Five-Factor Inventory; PGWBI = Psychological General Well-Being Index: AA = Absence of Anxiety; AD = Absence of Depression; PWB = Positive well-being; SC = Self-control; GH = General Health; VIT = Vitality; DERS = Difficulties in Emotion Regulation Scale; TAS-20 = The Toronto Alexithymia Scale-20: DIF = Difficulties Identifying Feelings; DDF = Difficulties Describing Feelings; EOT = Externally Oriented Thinking; RME-T = Reading the Mind in the Eyes-Test.

* Significant at $\alpha \leq .05$ after Holm-Bonferroni correction significance testing;

** Significant at $\alpha \leq .01$ after Holm-Bonferroni correction significance testing.

For Peer Review Only

ONLINE VS. PAPER-AND-PENCIL ITALIAN QCAE

Table 7. Structural Invariance of the QCAE (M1) between the Paper and Pencil and Online Administrations.

Model 1	χ^2	df	p	$\Delta\chi^2$	Δdf	p	AIC	RMSEA (90% CI)	SRMR	CFI	NNFI
1 Configural Invariance	526.7	160	<.001	–	–	–	686.7	0.08 (0.07 - 0.09)	0.08	0.95	0.93
2 Metric Invariance	538.1	170	<.001	11.4	10	0.33	678.1	0.08 (0.07 - 0.09)	0.08	0.95	0.93
3 Scalar Invariance	538.1	180	<.001	0.0	10	1.00	718.1	0.08 (0.07 - 0.08)	0.08	0.95	0.94
4 Invariant Unique Variance	565.7	195	<.001	27.6	15	0.02	715.7	0.07 (0.07 - 0.08)	0.08	0.95	0.94

AIC = Akaike's Information Criterion; RMSEA = root mean square error of approximation; CI = Confidence Interval; SRMR = standardized root mean square residual; CFI = compared fit index; NNFI = non-normed fit index.

For Peer Review Only

ONLINE VS. PAPER-AND-PENCIL ITALIAN QCAE

Table 8. Comparison between QCAE Scores from Paper-and-Pencil and Online Administrations.

	Paper-and-pencil Dataset		Online Dataset		*t*	*df*	*Uncorr. p*	*d*
	M	*SD*	*M*	*SD*				
Perspective Taking								
Men	28.0	5.1	29.7	4.6	-2.12	164	.04	-.34
Women	30.0	5.0	31.4	4.7	-3.21	522	<.01*	-.28
Entire Sample	29.5	5.1	31.0	4.7	-3.99	690	<.01**	-.31
Online Simulation								
Men	26.0	4.4	27.1	3.8	-1.60	164	.11	-.26
Women	27.3	4.8	28.0	4.2	-1.88	552	.06	-.17
Entire Sample	27.0	4.7	27.8	4.2	-2.53	690	.01	-.20
Emotion Contagion								
Men	9.7	2.6	9.4	2.5	.83	164	.41	.13
Women	10.9	2.4	11.4	2.7	-2.53	522	.01	-.22
Entire Sample	10.6	2.5	11.0	2.8	-2.16	690	.03	-.17
Proximal Responsivity								
Men	10.6	2.3	11.1	2.3	-1.35	164	.18	-.22
Women	12.3	2.1	12.8	2.2	-2.84	522	<.01	-.25
Entire Sample	11.9	2.3	12.5	2.3	-3.48	690	<.01*	-.27
Peripheral Responsivity								
Men	10.3	2.5	9.8	2.4	1.12	164	.26	.18
Women	12.0	2.2	12.4	2.5	-1.86	436.5[a]	.06	-.17
Entire Sample	11.5	2.4	11.8	2.7	-1.54	562.1[a]	.12	-.12
Cognitive Empathy								
Men	54.1	8.3	56.8	6.2	-2.27	164	.03	-.36
Women	57.3	8.3	59.4	7.5	-3.03	522	<.01*	-.27
Entire Sample	56.5	8.4	58.9	7.3	-3.90	690	<.01**	-.30
Affective Empathy								
Men	30.6	5.4	30.3	5.1	.33	164	.74	.05
Women	35.2	4.9	36.6	5.6	-3.17	444.7[a]	<.01*	-.29
Entire Sample	33.9	5.5	35.3	6.1	-3.02	567.8[a]	<.01*	-.24
Total Score								
Men	84.7	10.1	87.2	8.7	-1.61	164	.11	-.26
Women	92.4	11.0	96.1	10.7	-3.79	522	<.01**	-.33
Entire Sample	90.4	11.2	94.2	10.9	-4.38	690	<.01**	-.34

[a] Because homoscedasticity could not be assumed, the Welch–Satterthwaite method was used to adjust degrees of freedom

* Significant at $\alpha \leq .05$ after Holm-Bonferroni correction significance testing.

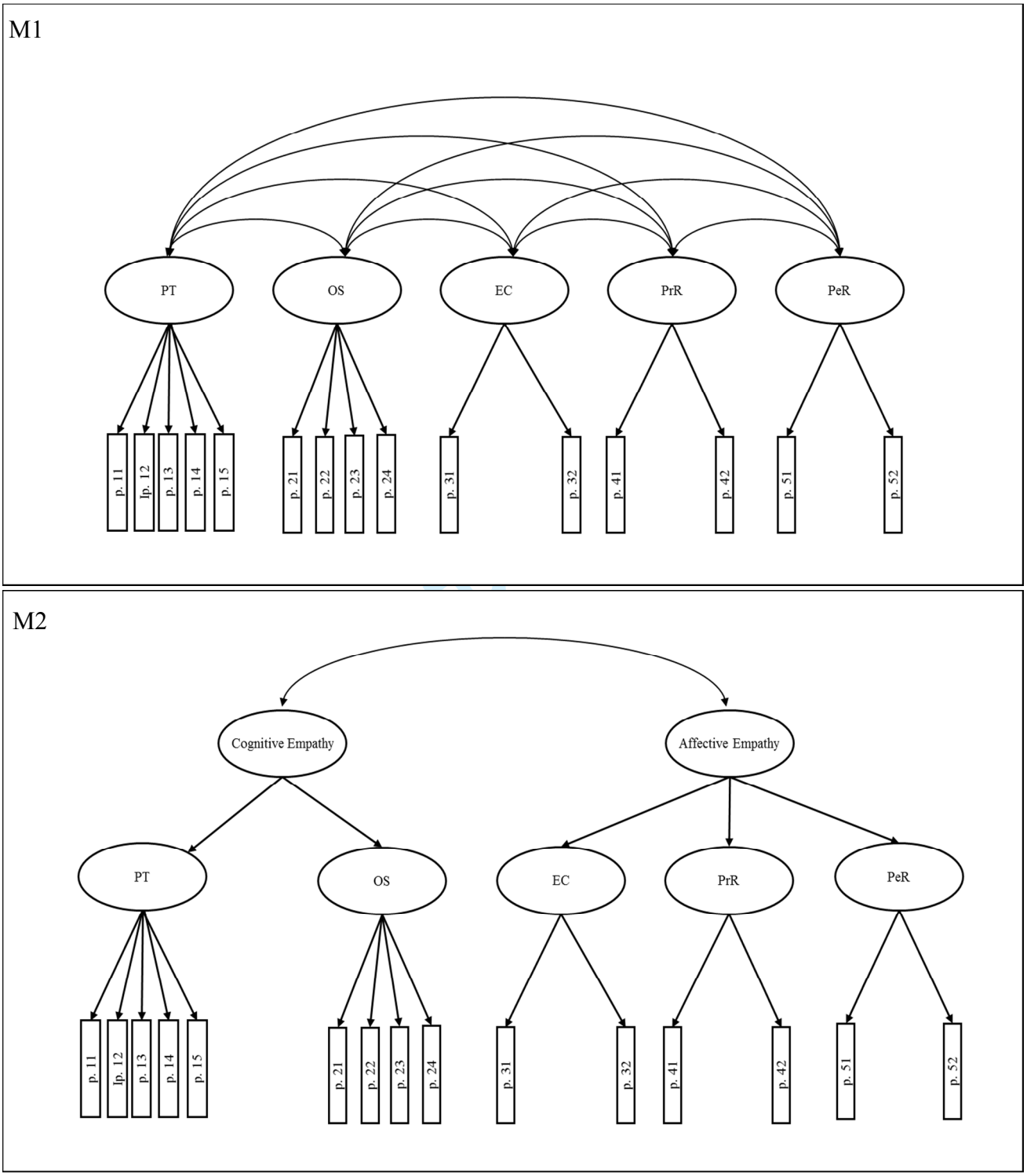
** Significant at $\alpha \leq .01$ after Holm-Bonferroni correction significance testing.

URL: http://mc.manuscriptcentral.com/JPersAssess Email: jpa_office@emich.edu

For Peer Review Only

ONLINE VS. PAPER-AND-PENCIL ITALIAN QCAE

Figure 1. Graphical Representation of the Inspected Factor Structure Models.



PT = Perspective Taking; OS = Online Simulation; EC = Emotion Contagion; PrR = Proximal Responsivity; PeR = Peripheral Responsivity; CE = Cognitive Empathy; AE = Affective Empathy. The labels "P 11, P 12, …, P 52" refer to the same item parcels utilized by Reniers et al. (2011).